Evaluación de Impacto de politicas Públicas
Clase Modelo
Tamaño de Muestra

Edinson Tolentino
email: edinson.tolentino@gmail.com

Twitter: @edutoleraymondi

ExpertiseMends Perú

7 de diciembre de 2025

Contenido

Tópicos

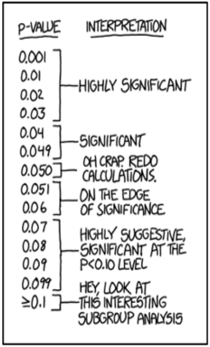
Sesiones	Tópicos	Descripccion
Sesion 1	**Regresión lineal**	Teoría y Práctica
	Regresión Lineal: estimaciones	
	Regresión Lineal: prediccion	
Sesion 2	**Calculo de Poder**	Teoría y Práctica
	Test de entrada (Inferencia)	
	Hipótesis nula y alternativa	
	Revisión de conceptos de muestras	
	Efectos de tamaños	
	Simulaciones: poder estadisticos y regresion lineal	
Sesion 3	**Diferencia en Diferencia**	Teoría y Práctica
	Test de entrada (Calculo de Poder)	
	Conceptos básicos	
	Estimación , inferencia y diagnostico para DiD	
Sesion 4	**Regresión Discontinua**	Teoría y Práctica
	Test de entrada (DiD)	
	RRD Sharp	
	Estimación , inferencia y diagnostico en RDD	
Sesion 5	**Inferencia de cluster y regresión lasso**	Teoría y Práctica
	Test de entrada (RDD)	
	Selección y predicción de modelo para Lasso	
	Inferencia para Lasso	
	Cómo afectan los clusters a la inferencia	
	El cluster robust variance estimator (CRVE)	
	Alternativas cuando los supuestos del CRVE no se cumplen	

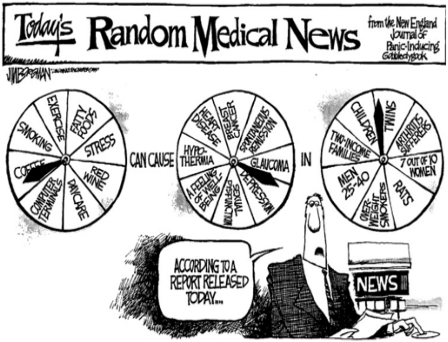
Cuadro: Topicos.

Logro de la sesión

- ✓ Conceptos de poder y tamaño de muestra
- ✓ Aplicación de calculo de poder
- ✓ Procedimiento para realizar simulaciones de calculo de poder estadistico

Introducción





Pregunta

- En 1970, la presión arterial sistólica (PAS) promedio en los Estados Unidos era de 100 mm/Hg.
- Debido al aumento en la prevalencia de la obesidad (que está asociada con una PAS más alta), podemos plantear que el PAS promedio en 2019 es de 110 mm/Hg.
- **Pregunta**: ¿Cómo podemos determinar si la PAS promedio en 2019 es 100 o 110, y cuántas personas debemos medir para responder esta pregunta?

Hipotesis Nula y Alternativa

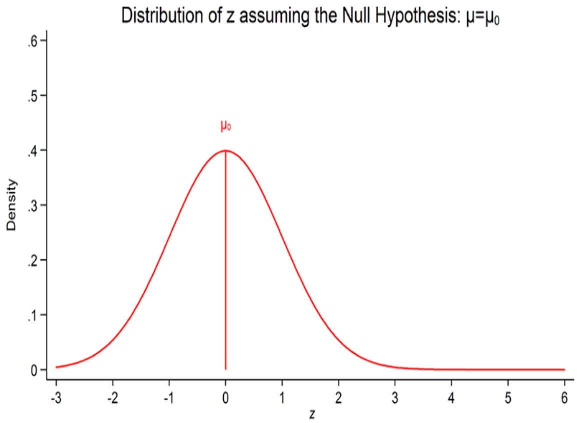
- Hipotesis Nula y Alternativa:
  1. Informalmente:
     1.1 Hipotesis Nula: el promedio de SBP en la población en 2019 es 100 mm/Hg
     1.2 Hipotesis Alternativa: el promedio de SBP en la población en 2019 es 110 mm/Hg
  2. Formalmente:
     2.1 Hipotesis Nula: La presión arterial sistólica (PAS) promedio de la población en 2019 tiene una distribución normal con media $\mu = 100$ mm/Hg y desviacion estandar $\sigma = 10$
     2.2 Hipotesis Alternativa: La presión arterial sistólica (PAS) promedio de la población en 2019 tiene una distribución normal con media $\mu = 110$ mm/Hg y desviacion estandar $\sigma = 10$
  3. Ecuaciones

$$\text{Hipotesis Nula} \qquad \mu_0 = \frac{100 - 100}{10} = 0$$

$$\text{Hipotesis Alternativa} \qquad \mu_A = \frac{110 - 100}{10} = 1$$

Hipotesis Nula y Alternativa

- Distribucion bajo la `hipotesis nula`



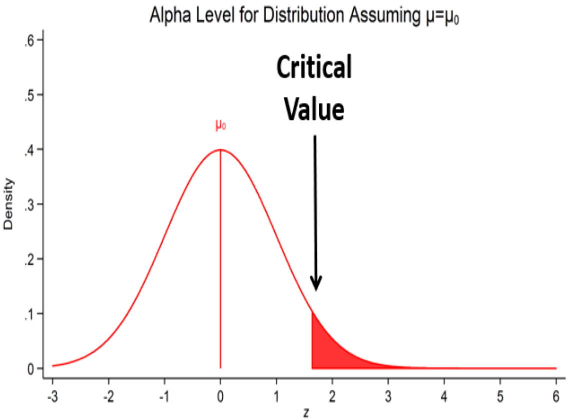
- Prueba (test) de una sola muestra promedio

$$z = \frac{\overline{x} - \mu_0}{\frac{\sigma}{\sqrt{n}}}$$

$$z = \frac{\overline{x} - 0}{\frac{1}{\sqrt{n}}}$$

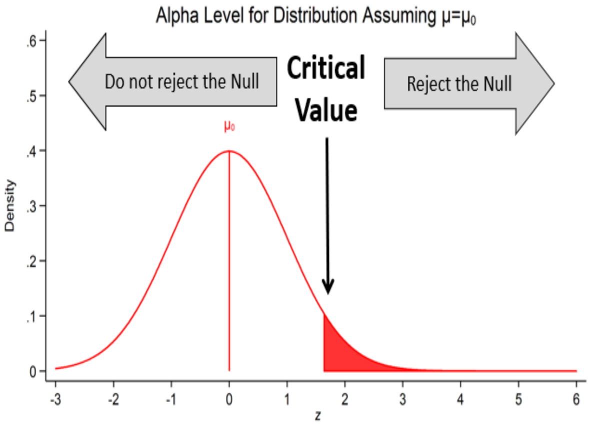
Hipotesis Nula y Alternativa

- Distribucion bajo la `hipotesis nula`



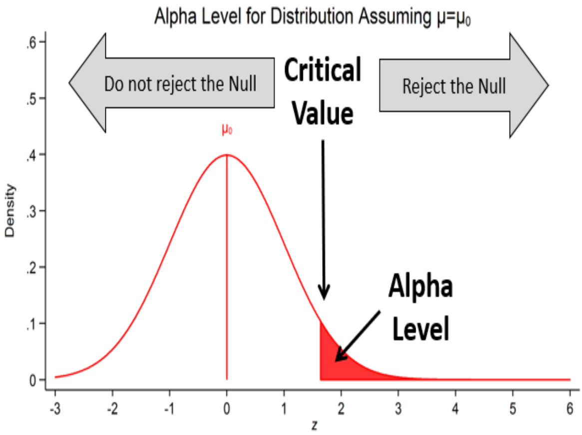
Valor critico

- Valor critico y nivel de alpha



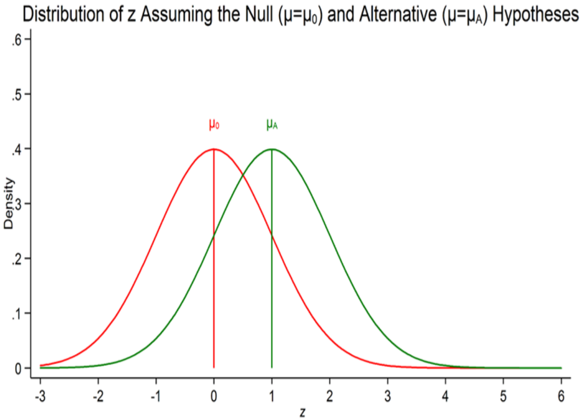
Valor critico

- Valor critico y nivel de alpha



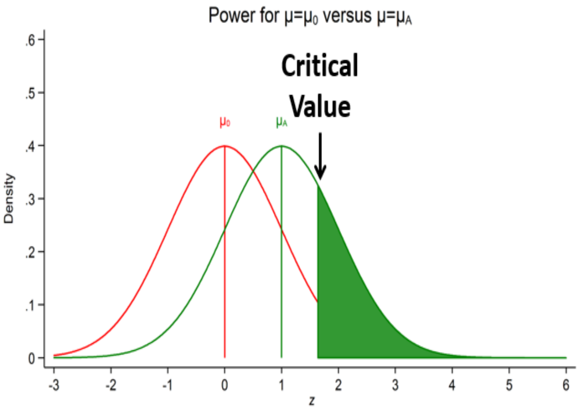
Valor critico

- Distribución bajo una hipotesis alternativa



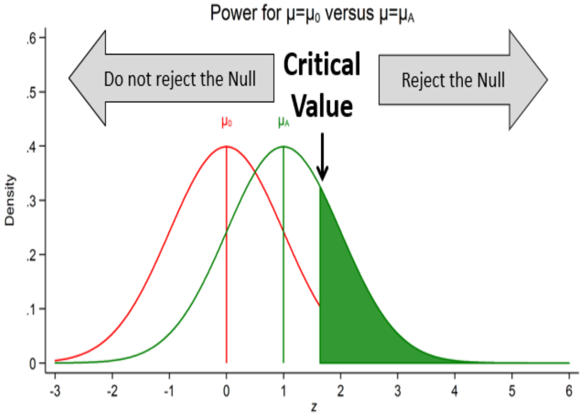
Poder

- Valor critico y poder



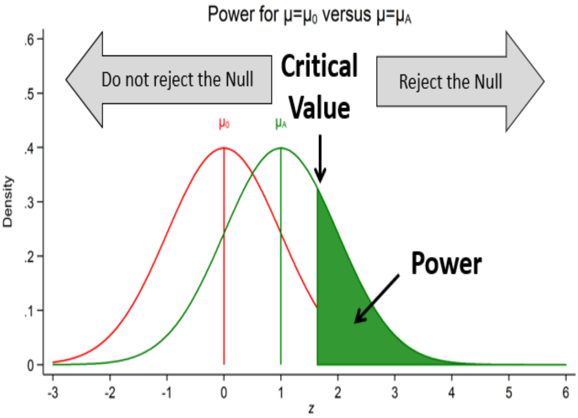
Poder

- Valor critico y poder



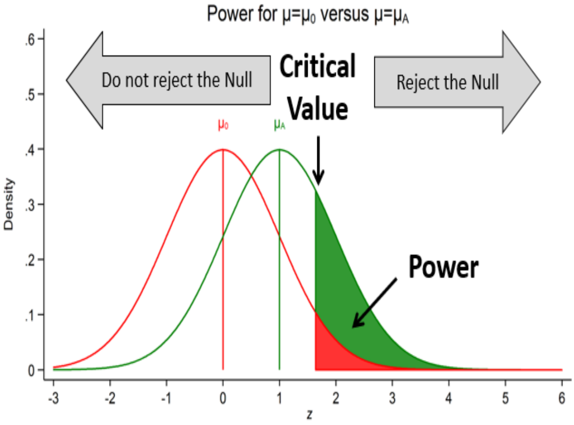
Poder

- Valor critico y poder



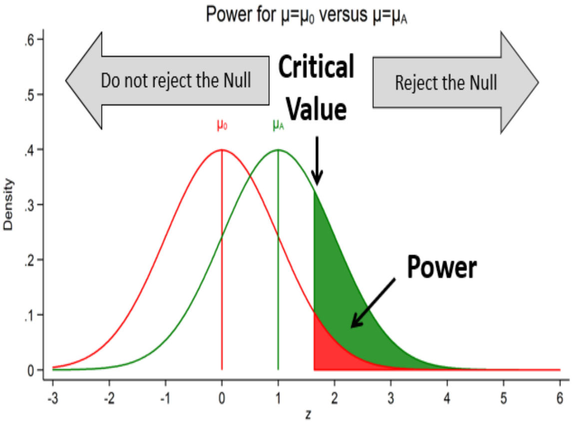
Poder etadistico

$$\text{Poder} = P\left(\text{z} > \text{Valor critico} \middle| \text{alternativa es verdadero}\right)$$



Poder y nivel de α

- **Nivel de alpha**: probabilidad de rechazar la hipotesis nula asumiendo que la hipotesis nula es verdadero
- **Poder estadistico**: probabilidad de rechazar la hipotesis nula asumiendo que la hipotesis alternativa es verdadero

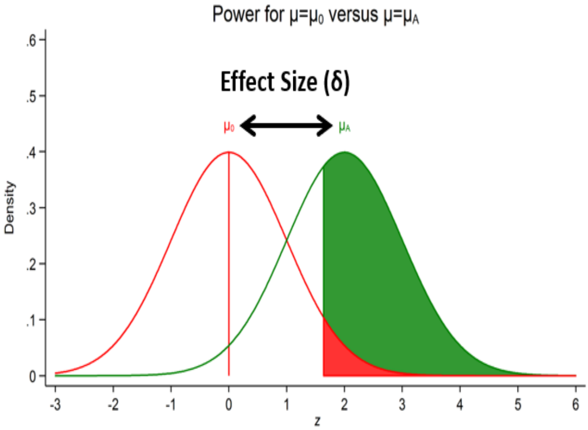


Efecto de tamaño

- **Efecto de tamaño**: el efecto de tamaño se puede definir de dos maneras:

No escalado $\delta = \mu_A - \mu_0$

Escalado $\delta = \dfrac{\mu_A - \mu_0}{\sigma}$

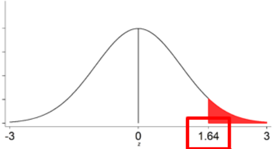


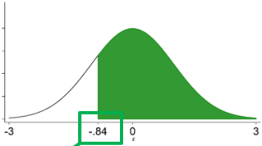
Efecto de tamaño

- Para resolver por tamaño de muestra:
  - **Alpha**
    - Una o dos colas
    - asignar α de 0.05 o 0.01
  - Poder
    - Establecido por tradicion
  - Efecto de tamaño
    - Diferencia entre los valores de las hipotesis de parametros
  - Varianza
    - Basados sobre estudios previos o informacion pilotos

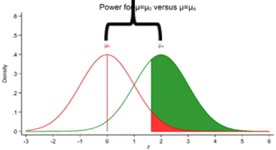
- Supuestos:
  1. Hipotesis nula: $\mu = \mu_0 = 0$
  2. Hipotesis alternativa: $\mu = \mu_A = 1$
  3. Varianza conocida: $\sigma^2 = 1$
  4. Una cola : $\alpha = 0.05$
  5. Poder : $1 - \beta = 0.80$
  6. Efecto de tamaño: $\delta = \left(\frac{\mu_A - \mu_0}{\sigma}\right) = 1$

Efecto de tamaño

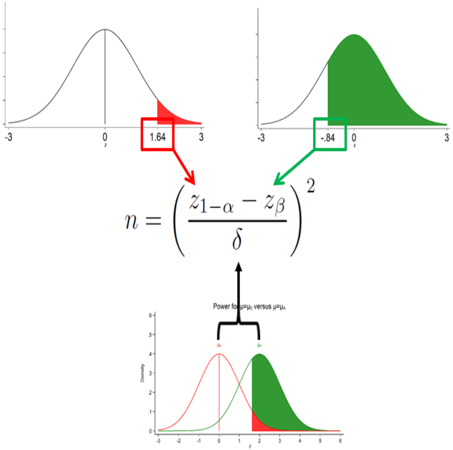




$$n = \left(\frac{z_{1-\alpha} - z_{\beta}}{\delta} \right)^2$$



tamaño muestra



- Supuestos:
  1. Hipotesis nula: $\mu = \mu_0 = 0$
  2. Hipotesis alternativa: $\mu = \mu_A = 1$
  3. Varianza conocida: $\sigma^2 = 1$
  4. Una cola : $\alpha = 0.05$
  5. Poder : $1 - \beta = 0.80$
  6. Efecto de tamaño: $\delta = \left(\frac{\mu_A - \mu_0}{\sigma}\right) = 1$
- Calculo de tamaño muestra

$$n = \left(\frac{z_{1-\alpha} - z_\beta}{\delta}\right)^2$$

$$n = \left(\frac{1.64 - (-0.80)}{1}\right)^2$$

$$n = 6.15$$

$$n = 7$$

tamaño muestra

```
. power onemean 0 1, knownsd onesided

Estimated sample size for a one-sample mean test
z test
Ho: m = m0  versus  Ha: m > m0

Study parameters:

        alpha =    0.0500
        power =    0.8000
        delta =    1.0000
           m0 =    0.0000
           ma =    1.0000
           sd =    1.0000

Estimated sample size:

            N =         7
```

- Supuestos:
  1. Hipotesis nula: $\mu = \mu_0 = 0$
  2. Hipotesis alternativa: $\mu = \mu_A = 1$
  3. Varianza conocida: $\sigma^2 = 1$
  4. Una cola : $\alpha = 0.05$
  5. Poder : $1 - \beta = 0.80$
  6. Efecto de tamaño: $\delta = \left(\frac{\mu_A - \mu_0}{\sigma}\right) = 1$
- Calculo de tamaño muestra

$$n = \left(\frac{z_{1-\alpha} - z_\beta}{\delta}\right)^2$$

$$n = \left(\frac{1.64 - (-0.80)}{1}\right)^2$$

$$n = 6.15$$

$$n = 7$$

Cálculo de poder

- Supuestos:
  1. Hipotesis nula: $\mu = \mu_0 = 0$
  2. Hipotesis alternativa: $\mu = \mu_A = 1$
  3. Varianza conocida: $\sigma^2 = 1$
  4. Una cola : $\alpha = 0.05$
  5. Tanaño de muestra: $n = 5$
  6. Efecto de tamaño: $\delta = \left(\frac{\mu_A - \mu_0}{\sigma}\right) = 1$

- Calculo de tamaño muestra

$$\pi = \Phi\left(\sqrt{n}\,\delta - z_{1-\alpha}\right)$$

$$\pi = \Phi\left(\sqrt{5}\,(1) - 1.64\right)$$

$$\pi = \Phi\,(0.5961)$$

$$\pi = 0.7244$$

Cálculo de poder

```
. power onemean 0 1, n(5) knownsd onesided

Estimated power for a one-sample mean test
z test
Ho: m = m0   versus   Ha: m > m0

Study parameters:

        alpha =     0.0500
            N =          5
        delta =     1.0000
           m0 =     0.0000
           ma =     1.0000
           sd =     1.0000

Estimated power:

        power =     0.7228
```

- Calculo de tamaño muestra

$$\pi = \Phi\left(\sqrt{n}\,\delta - z_{1-\alpha}\right)$$

$$\pi = \Phi\left(\sqrt{5}\ (1) - 1.64\right)$$

$$\pi = \Phi\left(0.5961\right)$$

$$\pi = 0.7244$$

Efecto de tamaño

- Supuestos:
  1. Hipotesis nula: $\mu = \mu_0 = 0$
  2. Hipotesis alternativa: $\mu = \mu_A = 1$
  3. Varianza conocida: $\sigma^2 = 1$
  4. Una cola : $\alpha = 0.05$
  5. Tamaño de muestra: $n = 5$
  6. Poder : $1 - \beta = 0.80$
  7. Varianza: basado en el conocimiento de data o muestra piloto

- Calculo de efecto de tamaño

$$|\delta| = \frac{(z_{1-\alpha} - z_{\beta})}{\sqrt{n}}$$

$$|\delta| = \frac{(1.64 - (-0.84))}{\sqrt{5}}$$

$$|\delta| = 1.11$$

Efecto de tamaño

```
. power onemean 0, n(5) power(0.8) knownsd onesided

Estimated target mean for a one-sample mean test
z test
Ho: m = m0  versus  Ha: m > m0; ma > m0

Study parameters:

        alpha =    0.0500
        power =    0.8000
            N =         5
           m0 =    0.0000
           sd =    1.0000

Estimated effect size and target mean:

        delta =    1.1120
           ma =    1.1120
```

- Calculo de efecto de tamaño

$$|\delta| = \frac{(z_{1-\alpha} - z_{\beta})}{\sqrt{n}}$$

$$|\delta| = \frac{(1.64 - (-0.84))}{\sqrt{5}}$$

$$|\delta| = 1.11$$

Resumen

- Para resolver el tamaño de muestra

$$n = \left(\frac{z_{1-\alpha} - z_{\beta}}{\delta} \right)^2$$

- Para resolver el poder

$$\pi = \Phi \left(\sqrt{n}\,\delta - z_{1-\alpha} \right)$$

- Paara poder resolver el efecto de tamaño

$$|\delta| = \frac{\left(z_{1-\alpha} - z_{\beta}\right)}{\sqrt{n}}$$

Tamaño muestra: RCT

- **Potencia o Poder:** probabilidad de que, para un tamaño de efecto dado y un nivel de significancia estadística dado, podamos rechazar la hipótesis de un efecto nulo.

Tamaño muestra: RCT

- **Potencia o Poder:** probabilidad de que, para un tamaño de efecto dado y un nivel de significancia estadística dado, podamos rechazar la hipótesis de un efecto nulo.
- El estimador de ATE en un experimento aletorizado simple es

Tamaño muestra: RCT

- **Potencia o Poder:** probabilidad de que, para un tamaño de efecto dado y un nivel de significancia estadística dado, podamos rechazar la hipótesis de un efecto nulo.
- El estimador de ATE en un experimento aletorizado simple es

$$\widehat{\beta} = \overline{Y}_1 - \overline{Y}_0$$

- Donde $\overline{Y}_k$ es el promedio del producto en el tratamiento ($k = 1$) or control ($k = 0$)

Tamaño muestra: RCT

- **Potencia o Poder:** probabilidad de que, para un tamaño de efecto dado y un nivel de significancia estadística dado, podamos rechazar la hipótesis de un efecto nulo.
- El estimador de ATE en un experimento aletorizado simple es

$$\widehat{\beta} = \overline{Y}_1 - \overline{Y}_0$$

- Donde $\overline{Y}_k$ es el promedio del producto en el tratamiento $(k = 1)$ or control $(k = 0)$
- El error estandar de $\widehat{\beta}$ es

$$se(\widehat{\beta}) = \frac{1}{P(1-P)} \frac{\sigma^2}{N}$$

- Donde σ^2 es la varianza. Si β_i es heterogeneo, entonces la varianza puede diferir.
- N es el tamaño de la muestra, y P es la proporcion asignada al grupo de tratamiento.

Ejercicio Númerico: tamaño muestral

- Supongamos que queremos ser capaces de detectar un efecto de al menos a con una potencia determinada $\kappa\,\%$

Ejercicio Númerico: tamaño muestral

- Supongamos que queremos ser capaces de detectar un efecto de al menos a con una potencia determinada $\kappa\,\%$
- Esto significa que, si el efecto verdadero es α, queremos elegir nuestro tamaño de muestra de modo que rechacemos la hipótesis nula de que el efecto es 0 con probabilidad $\kappa\,\%$ cuando el nivel de significancia es α
  - ✓ α es usualmente elegido por un $5\,\%$

Tamaño muestra: RCT

- Para lograr un poder de κ, debemos tener: $\beta > (t_{1-\kappa} + t_{\alpha}) \times se(\widehat{\beta})$

Tamaño muestra: RCT

- Para lograr un poder de κ, debemos tener: $\beta > (t_{1-\kappa} + t_{\alpha}) \times se(\widehat{\beta})$
- El tamaño mínimo del efecto detectable para una potencia dada,κ, un nivel de significancia α, un tamaño de muestra N y la proporción de sujetos asignados al grupo de tratamiento, P, está dado po

Tamaño muestra: RCT

- Para lograr un poder de κ, debemos tener: $\beta > (t_{1-\kappa} + t_{\alpha}) \times se(\widehat{\beta})$
- El tamaño mínimo del efecto detectable para una potencia dada,κ, un nivel de significancia α, un tamaño de muestra N y la proporción de sujetos asignados al grupo de tratamiento, P, está dado po

$$MDE = (t_{1-\kappa} + t_{\alpha}) \times \sqrt{\frac{1}{P(1-P)}}\sqrt{\frac{\sigma^2}{N}}$$

Tamaño muestra: RCT

- Para lograr un poder de κ, debemos tener: $\beta > (t_{1-\kappa} + t_{\alpha}) \times se(\widehat{\beta})$
- El tamaño mínimo del efecto detectable para una potencia dada,κ, un nivel de significancia α, un tamaño de muestra N y la proporción de sujetos asignados al grupo de tratamiento, P, está dado po

$$MDE = (t_{1-\kappa} + t_{\alpha}) \times \sqrt{\frac{1}{P(1-P)}}\sqrt{\frac{\sigma^2}{N}}$$

  - Esto es para una prueba de una sola cola. Para una prueba de dos colas, reemplace el valor correspondiente t_{α} con $t_{\frac{\alpha}{2}}$

Tamaño muestra: RCT

- Para lograr un poder de κ, debemos tener: $\beta > (t_{1-\kappa} + t_{\alpha}) \times se(\widehat{\beta})$
- El tamaño mínimo del efecto detectable para una potencia dada,κ, un nivel de significancia α, un tamaño de muestra N y la proporción de sujetos asignados al grupo de tratamiento, P, está dado po

$$MDE = (t_{1-\kappa} + t_{\alpha}) \times \sqrt{\frac{1}{P(1-P)}}\sqrt{\frac{\sigma^2}{N}}$$

  - Esto es para una prueba de una sola cola. Para una prueba de dos colas, reemplace el valor correspondiente t_{α} con $t_{\frac{\alpha}{2}}$
  - La fórmula puede usarse tanto para calcular el tamaño mínimo detectable del efecto (MDE) dado un tamaño de muestra, como para calcular el tamaño de muestra dado un MDE

Tamaño muestra: RCT

- Para lograr un poder de κ, debemos tener: $\beta > (t_{1-\kappa} + t_{\alpha}) \times se(\widehat{\beta})$
- El tamaño mínimo del efecto detectable para una potencia dada,κ, un nivel de significancia α, un tamaño de muestra N y la proporción de sujetos asignados al grupo de tratamiento, P, está dado po

$$MDE = (t_{1-\kappa} + t_{\alpha}) \times \sqrt{\frac{1}{P(1-P)}}\sqrt{\frac{\sigma^2}{N}}$$

  - Esto es para una prueba de una sola cola. Para una prueba de dos colas, reemplace el valor correspondiente t_{α} con $t_{\frac{\alpha}{2}}$
  - La fórmula puede usarse tanto para calcular el tamaño mínimo detectable del efecto (MDE) dado un tamaño de muestra, como para calcular el tamaño de muestra dado un MDE
- Existe un trade-off entre el poder (potencia) y tamaño

Tamaño muestra: RCT

- **Intervención:** Ofrecer un subsidio escolar a niños de 12 a 16 años condicionado a la asistencia.

Tamaño muestra: RCT

- **Intervención:** Ofrecer un subsidio escolar a niños de 12 a 16 años condicionado a la asistencia.
- **Variable de resultado:** Asistencia escolar de los niños en el grupo elegible y de los más pequeños.

Tamaño muestra: RCT

- **Intervención:** Ofrecer un subsidio escolar a niños de 12 a 16 años condicionado a la asistencia.
- **Variable de resultado:** Asistencia escolar de los niños en el grupo elegible y de los más pequeños.
  - ✓ El efecto probablemente sea positivo para los niños elegibles, pero podría ser cualquiera para los más pequeños.

Tamaño muestral

- Los responsables de políticas han decidido que un efecto de 0.1 o 10 % haría que el programa valiera la pena.

Tamaño muestral

- Los responsables de políticas han decidido que un efecto de 0.1 o 10 % haría que el programa valiera la pena.
- Esto implica que poder proceder al calculo de poder
- Supongamos que la proporción que asiste antes de la intervención es 0.4. Esto implica que

$$\frac{\sigma^2}{N} = \frac{\Pr(1-\Pr)}{N} = \frac{0.24}{N}$$

Tamaño muestral

- Los responsables de políticas han decidido que un efecto de 0.1 o 10 % haría que el programa valiera la pena.
- Esto implica que poder proceder al calculo de poder
- Supongamos que la proporción que asiste antes de la intervención es 0.4. Esto implica que

$$\frac{\sigma^2}{N} = \frac{\Pr(1-\Pr)}{N} = \frac{0.24}{N}$$

- Queremos seleccionar el tamaño de muestra total, N, de manera que $P = \frac{1}{2}$.

Tamaño muestral

- Los responsables de políticas han decidido que un efecto de 0.1 o 10 % haría que el programa valiera la pena.
- Esto implica que poder proceder al calculo de poder
- Supongamos que la proporción que asiste antes de la intervención es 0.4. Esto implica que

$$\frac{\sigma^2}{N} = \frac{\mathrm{Pr}(1 - \mathrm{Pr})}{N} = \frac{0.24}{N}$$

- Queremos seleccionar el tamaño de muestra total, N, de manera que $P = \frac{1}{2}$.
- Además, fijamos $\alpha = 5\,\%$, y deseamos rechazar la hipótesis nula de ausencia de efecto con al menos $\kappa = 0.8$ si el efecto verdadero es mayor que 0.1.Entonces:

$$(t_{1-\kappa} + t_{\alpha}) = 0.84 + 1.96 = 2.8$$

Tamaño muestral

- Los responsables de políticas han decidido que un efecto de 0.1 o 10 % haría que el programa valiera la pena.
- Esto implica que poder proceder al calculo de poder
- Supongamos que la proporción que asiste antes de la intervención es 0.4. Esto implica que

$$\frac{\sigma^2}{N} = \frac{\Pr(1-\Pr)}{N} = \frac{0.24}{N}$$

- Queremos seleccionar el tamaño de muestra total, N, de manera que $P = \frac{1}{2}$.
- Además, fijamos $\alpha = 5\,\%$, y deseamos rechazar la hipótesis nula de ausencia de efecto con al menos $\kappa = 0.8$ si el efecto verdadero es mayor que 0.1.Entonces:

$$(t_{1-\kappa} + t_\alpha) = 0.84 + 1.96 = 2.8$$

$$MDE = (t_{1-\kappa} + t_\alpha) \times \sqrt{\frac{1}{P(1-P)}}\sqrt{\frac{\sigma^2}{N}}$$

Tamaño muestral

- Los responsables de políticas han decidido que un efecto de 0.1 o 10 % haría que el programa valiera la pena.
- Esto implica que poder proceder al calculo de poder
- Supongamos que la proporción que asiste antes de la intervención es 0.4. Esto implica que

$$\frac{\sigma^2}{N} = \frac{\Pr(1-\Pr)}{N} = \frac{0.24}{N}$$

- Queremos seleccionar el tamaño de muestra total, N, de manera que $P = \frac{1}{2}$.
- Además, fijamos $\alpha = 5\,\%$, y deseamos rechazar la hipótesis nula de ausencia de efecto con al menos $\kappa = 0.8$ si el efecto verdadero es mayor que 0.1.Entonces:

$$(t_{1-\kappa} + t_\alpha) = 0.84 + 1.96 = 2.8$$

$$MDE = (t_{1-\kappa} + t_\alpha) \times \sqrt{\frac{1}{P(1-P)}}\sqrt{\frac{\sigma^2}{N}}$$

$$0.1 = 2.8 \times \sqrt{\frac{1}{0.5(1-0.5)}}\sqrt{\frac{0.24}{N}} \equiv 753$$

Tamaño muestral

- Los responsables de políticas han decidido que un efecto de 0.1 o 10 % haría que el programa valiera la pena.
- Esto implica que poder proceder al calculo de poder
- Supongamos que la proporción que asiste antes de la intervención es 0.4. Esto implica que

$$\frac{\sigma^2}{N} = \frac{\Pr(1-\Pr)}{N} = \frac{0.24}{N}$$

- Queremos seleccionar el tamaño de muestra total, N, de manera que $P = \frac{1}{2}$.
- Además, fijamos $\alpha = 5\,\%$, y deseamos rechazar la hipótesis nula de ausencia de efecto con al menos $\kappa = 0.8$ si el efecto verdadero es mayor que 0.1.Entonces:

$$(t_{1-\kappa} + t_{\alpha}) = 0.84 + 1.96 = 2.8$$

$$MDE = (t_{1-\kappa} + t_{\alpha}) \times \sqrt{\frac{1}{P(1-P)}}\sqrt{\frac{\sigma^2}{N}}$$

$$0.1 = 2.8 \times \sqrt{\frac{1}{0.5(1-0.5)}}\sqrt{\frac{0.24}{N}} \equiv 753$$

- Entonces, tenemos aproximadamente 377 individuos en cada grupo

Como usar program

- Pasos:

- La potencia es la proporción de veces que se rechaza la hipótesis nula.

Como usar program

- Pasos:
  1. Crear un conjunto de datos seudoaleatorio asumiendo que la hipótesis alternativa es verdadera.

- La potencia es la proporción de veces que se rechaza la hipótesis nula.

Como usar program

- Pasos:
  1. Crear un conjunto de datos seudoaleatorio asumiendo que la hipótesis alternativa es verdadera.
  2. Probar la hipótesis nula usando ese conjunto de datos.

- La potencia es la proporción de veces que se rechaza la hipótesis nula.

Como usar program

- Pasos:
  1. Crear un conjunto de datos seudoaleatorio asumiendo que la hipótesis alternativa es verdadera.
  2. Probar la hipótesis nula usando ese conjunto de datos.
  3. Guardar los resultados de la prueba de hipótesis (por ejemplo, “1” si se rechaza y “0” si no se rechaza).

- La potencia es la proporción de veces que se rechaza la hipótesis nula.

Como usar program

- Pasos:
  1. Crear un conjunto de datos seudoaleatorio asumiendo que la hipótesis alternativa es verdadera.
  2. Probar la hipótesis nula usando ese conjunto de datos.
  3. Guardar los resultados de la prueba de hipótesis (por ejemplo, "1" si se rechaza y "0" si no se rechaza).
  4. Repetir los Pasos 1 y 3 muchas veces (1000+).
- La potencia es la proporción de veces que se rechaza la hipótesis nula.

Como usar program

- Pasos:
  1. Crear un conjunto de datos seudoaleatorio asumiendo que la hipótesis alternativa es verdadera.
  2. Probar la hipótesis nula usando ese conjunto de datos.
  3. Guardar los resultados de la prueba de hipótesis (por ejemplo, "1" si se rechaza y "0" si no se rechaza).
  4. Repetir los Pasos 1 y 3 muchas veces (1000+).
- La potencia es la proporción de veces que se rechaza la hipótesis nula.
- Instrumentos:

Como usar program

- Pasos:
  1. Crear un conjunto de datos seudoaleatorio asumiendo que la hipótesis alternativa es verdadera.
  2. Probar la hipótesis nula usando ese conjunto de datos.
  3. Guardar los resultados de la prueba de hipótesis (por ejemplo, "1" si se rechaza y "0" si no se rechaza).
  4. Repetir los Pasos 1 y 3 muchas veces (1000+).
- La potencia es la proporción de veces que se rechaza la hipótesis nula.
- Instrumentos:
  - Uso de locales y macros

Como usar program

- Pasos:
  1. Crear un conjunto de datos seudoaleatorio asumiendo que la hipótesis alternativa es verdadera.
  2. Probar la hipótesis nula usando ese conjunto de datos.
  3. Guardar los resultados de la prueba de hipótesis (por ejemplo, “1” si se rechaza y “0” si no se rechaza).
  4. Repetir los Pasos 1 y 3 muchas veces (1000+).
- La potencia es la proporción de veces que se rechaza la hipótesis nula.
- Instrumentos:
  - Uso de locales y macros
  - Creaar pseudo-random datasets

Como usar program

- Pasos:
  1. Crear un conjunto de datos seudoaleatorio asumiendo que la hipótesis alternativa es verdadera.
  2. Probar la hipótesis nula usando ese conjunto de datos.
  3. Guardar los resultados de la prueba de hipótesis (por ejemplo, “1” si se rechaza y “0” si no se rechaza).
  4. Repetir los Pasos 1 y 3 muchas veces (1000+).
- La potencia es la proporción de veces que se rechaza la hipótesis nula.
- Instrumentos:
  - Uso de locales y macros
  - Creaar pseudo-random datasets
  - Como usar el programapara crear un simple comando

Como usar program

- En la simulación, ejecutarás repetidamente los comandos que generan datos aleatorios y prueban la hipótesis nula.

Como usar program

- En la simulación, ejecutarás repetidamente los comandos que generan datos aleatorios y prueban la hipótesis nula.
- Debes definir los parámetros de entrada de la simulación y devolver los resultados de las pruebas.

Como usar program

- En la simulación, ejecutarás repetidamente los comandos que generan datos aleatorios y prueban la hipótesis nula.
- Debes definir los parámetros de entrada de la simulación y devolver los resultados de las pruebas.
- Para hacerlo de forma eficiente, puedes crear tu propio comando en Stata usando program

Como usar program

- En la simulación, ejecutarás repetidamente los comandos que generan datos aleatorios y prueban la hipótesis nula.
- Debes definir los parámetros de entrada de la simulación y devolver los resultados de las pruebas.
- Para hacerlo de forma eficiente, puedes crear tu propio comando en Stata usando program
- El ejemplo dado define un comando `myprogram` que recibe el parametro `n()`, muestra su valor y lo devuelve.

Como usar program

- En la simulación, ejecutarás repetidamente los comandos que generan datos aleatorios y prueban la hipótesis nula.
- Debes definir los parámetros de entrada de la simulación y devolver los resultados de las pruebas.
- Para hacerlo de forma eficiente, puedes crear tu propio comando en Stata usando program
- El ejemplo dado define un comando `myprogram` que recibe el parametro `n()`, muestra su valor y lo devuelve.

```
capture program drop myprogram
program myprogram, rclass
        version 15.1
        syntax, n(integer)
        display "n = `n'"
        return scalar N = `n'
end
```

Como usar program

- Se define un programa llamado **simttest** que genera un conjunto de datos aleatorio según parámetros de entrada.

Como usar program

- Se define un programa llamado **simttest** que genera un conjunto de datos aleatorio según parámetros de entrada.
- El programa realiza una prueba de la hipótesis nula usando esos datos.

Como usar program

- Se define un programa llamado **simttest** que genera un conjunto de datos aleatorio según parámetros de entrada.
- El programa realiza una prueba de la hipótesis nula usando esos datos.
- Luego devuelve los resultados de dicha prueba; el bloque de código mostrado solo contiene comentarios.

Como usar program

- Se define un programa llamado **simttest** que genera un conjunto de datos aleatorio según parámetros de entrada.
- El programa realiza una prueba de la hipótesis nula usando esos datos.
- Luego devuelve los resultados de dicha prueba; el bloque de código mostrado solo contiene comentarios.

```
capture program drop myprogram
program myprogram, rclass
        version 15.1
        syntax, n(integer)
        display "n = `n'"
        return scalar N = `n'
end
```

Como usar program

- Mi programa

```
capture program drop simttest
program simttest, rclass
    version 15.1

    // DEFINE THE INPUT PARAMETERS AND THEIR DEFAULT VALUES
    syntax, n(integer)              ///  Sample size
          [ alpha(real 0.05)        ///  Alpha level
            m0(real 0)              ///  Mean under the null
            ma(real 1)              ///  Mean under the alternative
            sd(real 1) ]            //   Standard deviation

    // GENERATE THE RANDOM DATA AND TEST THE NULL HYPOTHESIS
    drawnorm y, n(`n') means(`ma') sds(`sd') clear
    ttest y = `m0'

    // RETURN RESULTS
    return scalar reject = (r(p)<`alpha')
end
```

- Simulaciones

```
. simttest, n(100) m0(70) ma(75) sd(15) alpha(0.05)
(obs 100)

One-sample t test
------------------------------------------------------------------------------
Variable |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
       y |     100    77.22324    1.597794    15.97794    74.05287    80.39361
------------------------------------------------------------------------------
    mean = mean(y)                                                t =   4.5208
Ho: mean = 70                                    degrees of freedom =       99
```

Como usar program

- El programa **simttest** realiza todo lo necesario para una iteración de la simulación.
- Luego se necesita una forma de ejecutar simttest muchas veces.
- También es necesario recopilar los resultados de cada iteración.
- El bloque de código muestra cómo usar simulate para lograr ambas tareas.

```
. simulate reject=r(reject), reps(100) seed(12345):        ///
          simttest, n(100) m0(70) ma(75) sd(15) alpha(0.05)

      command:  simttest, n(100) m0(70) ma(75) sd(15) alpha(0.05)
       reject:  r(reject)

Simulations (100)
----+--- 1 ---+--- 2 ---+--- 3 ---+--- 4 ---+--- 5
..................................................    50
..................................................   100
```

- Puedes usar summarize para calcular la media de reject, que equivale a la proporción de veces (de 100 iteraciones) que se rechazó la hipótesis nula. ¡Esa proporción es tu estimación del poder estadístico dados los parámetros de entrada!

```
. summarize reject

    Variable |        Obs        Mean    Std. Dev.       Min        Max
-------------+---------------------------------------------------------
      reject |        100         .91    .2876235          0          1
```

- La estimacion del calculo de poder es 91 %

Como usar program

```
capture program drop power_cmd_simttest
program power_cmd_simttest, rclass
    version 15.1

    // DEFINE THE INPUT PARAMETERS AND THEIR DEFAULT VALUES
    syntax, n(integer)          ///  Sample size
          [ alpha(real 0.05)    ///  Alpha level (default is 0.05)
            m0(real 0)          ///  Mean under the null (default is 0)
            ma(real 1)          ///  Mean under the alternative (default is 1)
            sd(real 1)          ///  Standard deviation (default is 1)
            reps(integer 100)]  //   Number of repetitions (default is 100)

    // GENERATE THE RANDOM DATA AND TEST THE NULL HYPOTHESIS
    quietly simulate reject=r(reject), reps(`reps'):        ///
                    simttest, n(`n') m0(`m0') ma(`ma') sd(`sd') alpha(`alpha')
    quietly summarize reject

    // RETURN RESULTS
    return scalar power = r(mean)
    return scalar N     = `n'
    return scalar alpha = `alpha'
    return scalar m0    = `m0'
    return scalar ma    = `ma'
    return scalar sd    = `sd'
end
```

Como usar program

```
. power simttest, n(100) m0(70) ma(75) sd(15)

Estimated power
Two-sided test
```

alpha	power	N	m0	ma	sd
.05	.91	100	70	75	15

```
. power simttest, n(50(10)100) m0(70) ma(75) sd(15) table graph

Estimated power
Two-sided test
```

alpha	power	N	m0	ma	sd
.05	.67	50	70	75	15
.05	.73	60	70	75	15
.05	.73	70	70	75	15
.05	.77	80	70	75	15
.05	.85	90	70	75	15
.05	.91	100	70	75	15